CHAPTER 1

INTRODUCTION

1.1 FOREWORD TO THE PRESENT STUDY

As the recently introduced GCSE examination system includes marks for practical skills in science subjects, it has become more important than ever before to learn how to teach these skills effectively.

Teaching of practical skills can be described as a success, provided the pupils can reproduce these skills accurately in similar circumstances, or use this knowledge to carry out similar operations in the future.

Accurate assessment, like effective teaching, could be an equally difficult task in which fairness and reliability are important criteria. The outcome of an assessment not only provides information about the standard achieved by the pupil, it also creates scope for further teaching plans.

Furthermore, assessment could be based either (i) on the observation and marking of the performance of the manipulative skills/techniques during the practical session (referred to as Method -T), or (ii) on the outcome of an extended task and marking of the final results/products submitted as a written report of the work (referred to as Method -P). As the execution of the manipulative skill is likely to affect the final result, assessment from either direction could be expected to be equally acceptable for grading the pupil, but it is perhaps worthwhile beforehand to compare data from both the sources and ascertain if a significant correlation exists between them.

As the correct assessment of manipulative skills during a practical session is a difficult task, because a teacher may have to judge a large number of pupils in a short period of time, and he may have preconceived views about the pupil's ability. As a result, it has been thought to be worthwhile to check the reliability of such assessment among a number of judges by comparing the marks given by them for a pupil's ability to execute certain manipulative skills.

1.2 Historical development of assessment in school science practical work

The following is a summary of a paper by Lock[1] in which he compiled in detail a history of the practical work in school science and its assessment.

Practical work in school science in Britain started in the second half of the nineteenth century. In the early stages practical work filled a largely supportive role, i.e. confirming the theory which has already been taught. A similar verification approach also existed in the USA.
It is towards the end of the nineteenth century, when Armstrong's heurism began to exert an influence on the nature of the practical work[2, 3, 4], that such a "discovery approach" to science had widespread effects on both sides of the Atlantic.

Following the First World war, the Thompson Committee[5] were critical of school practical work as a lot of time was devoted to laboratory work involving repetitive practical exercises.

The 1944 Education Act had little direct impact on practical work, but it did make recommendations as to the number and size of science rooms that a school should have[6].

The debate over effectiveness of individual practical work as compared with demonstrations went on from 1920s to the 1950s, particularly in the USA, and although the 1950s in America saw the planning of new courses, which heralded the return of discovery learning, controversy over this has continued[7, 8].

Meanwhile in Britain, the late 1940s and 1950s saw a continuation of the confirmatory role of practical work --- with demonstration moving especially up to the end of compulsory schooling (year 5). Study in the sixth form, featured more involvement of pupils in practical work, but it remained corroborative of theory or biased towards developing a high proficiency in a restrictive range of techniques, such as titration and dissection.

With practical work biased towards learning technique, it was not difficult to assess the skills in a three-hour practical examination. Such a situation persisted throughout the 1960s and into the 1980s. The influence of the Nuffield A-level Curriculum was, however, felt more broadly in the 1970s, with some A-level syllabuses changing the practical examination format and adopting internal teacher assessment of practical skills.

There were no practical examinations or internal assessment of practical skills with the Nuffield GCE O-level courses (although Nuffield O-level chemistry did provide internally assessed options)[9].

The other significant event of the 1960s, with implications for school practical work, was the introduction of the Certificate of Secondary Education (CSE). The teachers had the freedom of designing their own syllabuses and adopting their own method of assessment.

The extent to which continued assessment by internal means was employed in science syllabuses is well illustrated by Eggleston[10], and it is interesting to note that the following years saw a decline, at CSE level, of the compulsory practical examinations[11]. The other method of assessment that gained in popularity at CSE level was the project. This was clearly an adoption of ideas introduced by some of the Nuffield Advanced level schemes.

An event in the late 1970s with potential implications for practical work in schools and its assessment, was the speech of the then Prime Minister, James Callaghan, at Ruskin College in Oxford. This led to the establishment of APU (Assessment Performance Unit) whose framework included the assessment of practical skills[12].

In 1982, the Oxford Certificate of Educational Achievement (OCEA) announced provisions for graded tests which would incorporate assessment of practical skills[13].

When the national criteria for GCSE (General Certificate for Secondary Education) examinations in science were published in May 1985, they made it clear that[14]: "All schemes of assessment must allocate not less than 20% of the total marks to experimental and theoretical skills".

In summarising, it can be said that the relatively recent introduction of teacher-assessed techniques was influenced by Nuffield A-level courses and the CSE scheme, both of which involved the teacher in the role of assessor.

Development of GCSE has created the need for the internal assessment of practical skills of all candidates entered for science examinations at the age of 16

1.3 Review of some previous work on assessment of practical skills in school science.

Due to the growing important of science in the school curriculum and the increasing emphasis given in the assessment of practical work, the laboratory in some instances has been described as[15] - "...not just a place for demonstration and confirmation, but rather as the core of the Science learning process".

Until the late 1970s, in most cognitive assessment activity in science, importance was given to the recall of knowledge. At this time, Bloom's "Cognitive" taxonomy[16] was adopted as a basis for the assessment objective. However, in the 1980s, more importance has started to be given to the assessment of intellectual (cognitive) and practical (psychomotor) skills and abilities.

Abolition of the norm-referenced 2 tier examination system, viz. GCE and CSE and introduction of the criterion-referenced GCSE in their place, created a new situation in science education in schools. In this examination system, there is a compulsory element of teacher-assessed practical work, carried out by the pupils throughout the 2-year schooling period.

Some of the research work carried out in conjunction with various assessment projects (e.g. SAPA and NAEP in the USA, Nuffield Science 5-13, TVEI and GCSE in the UK) has been reviewed recently by Johnson[17].

Holstein and Lunetta[18], having reviewed a selection of past research work, suggested that the area of practical skills assessment has remained a neglected area. Questions ought to be raised (1) to know the nature of the work pupils are doing in the laboratory, and (2) to find an appropriate method of assessing their work. In this context the work of Olsen[19] can be referred to when making a clear distinction between development of knowledge (know-that) and development of skill (know-how). He stressed the fact that often skills are undervalued by using them as a means to gain knowledge, although acquisition of scientific skills is much more important than that. Similar stress was given to scientific skills by the DES/APU project[20], by Bryce et. al[21], who suggested the presence of evidence for the fact that pupils who may not do well in traditional assessment, giving emphasis on recall of knowledge, may demonstrate better performance in carrying out work involving scientific skills.

Regarding the in-school assessment of pupils' practical skills, the work of the TAPS group[21-26] is very familiar to many teachers. The TAPS project has evaluated techniques for the assessment of pupils' practical skills and produced guidance materials for teachers. The TAPS assessment pack for teachers has been produced to cover a large number of skills, viz. measuring, observing, recording, drawing inferences and problem solving etc.

However, a contradictory view is proposed by Miller[27] regarding the school science practical work. It has been suggested that the purpose of the experiment in a science laboratory will be to develop knowledge and understanding through interpretation, which is closer to the real science. This is very different from the traditional view of teaching various skills as has been advocated by the Nuffield reports[28] and then echoed by the DES reports[29]. The Nuffield report states that "pupils acquire the feeling of doing science, of being ...." a scientist for the day" '. According to the DES reports, "the essential characteristictic of education in science is that it introduces pupils to the methods of science".

Recently Woolnough and Toh[30], in reporting their work on assessment of practical work in the school laboratory, rejected the idea of assessing individual skills, i.e. "scientific processes" involved in various steps of the work. Instead they advocated the use of a "holistic approach", i.e.

the assessment of the final outcome of the whole experiment.

From this it appears that they have ignored the fact that science is a practically based subject and how efficiently the manipulative skills in practical work are performed is important in order to achieve correct results submitted in the form of a written report. In an experiment, a number of manipulative skills are involved. All of them may be important, but some are likely to be more important than others as their impact on the final result of the experiment is concerned. Furthermore, all these skills involved in an experiment cannot be assessed in one session because of various constraints. Performance of the practical skills and answering the check-list items may be complementary to each other, when a pupil is assessed for his/her practical work. However, either of these two criteria can be used as a basis for assessment, provided it is proved that there is a significant correlation between them. Woolnough and Toh[30] have reported a strong correlation between two methods of assessment and they tried to show that the final report is a good substitute of assessment in the laboratory. Perhaps it would be worthwhile to carry out a further study and check if there is a strong correlation between the assessment of the performance of skills at the work-place and assessment of the results reported on paper on the basis of their practical work.

In dismissing the idea of assessing practical skills, Woolnough and Toh[30] have cited the work of Miller and Driver[31], who share the view that the learning of an individual psychomotor skill by pupils in a laboratory is of no importance regarding their acquisition of scientific knowledge.

Miller and Driver[31] propose that "school science" is different from "real science", although they do not clearly define what the "real science" is or how it comes about. They failed to understand that various scientific skills learned by a pupil, and knowledge gained thereby in a school laboratory, enables him/her to perform a more complicated task in a scientific laboratory after leaving school. Furthermore these individual tasks taught and assessed in a school science laboratory are not seen as isolated process skills by the pupils. They aim to achieve a final objective for which various tasks are to be performed and in the end these activities would lead to a final result.

In their paper Miller and Driver[31] criticised the idea proposed by Bryce et. al.[21-26] who gave a considerable importance to the process skills.

Bryce and Robertson[32] replied to this criticism by saying - "Miller and Driver have taken restricted views of process in scientific, psychological and pedagogical terms and thus misrepresented the thinking behind the assessment materials developed by the TAPS research group[21-26]".

Pupils' understanding of science through reasoning and problem solving (discovery method) may be the final goal, but to achieve those, I feel that they should be taught and assessed on some basic scientific skills which will enable the pupils to carry out a more complicated task in their future life (perhaps as a professional scientist).

Although Miller and Driver[31] tried to relegate process skills to pedagogic means rather than to educational ends, they hope ultimately that "a critical appreciation of the way scientists work" will empower pupils to demonstrate their skills, which they consider desirable.

Bryce and Robertson[32] argued at length to counteract the criticisms put forward by Miller and Driver[31] against the process skills in the school science laboratory. The TAPS materials developed by Bryce et. al.[21-26] were not content-independent process skills. Pupils taught to follow the check-list of these skills will certainly have an opportunity to make inferences in scientific context at the end of the work.

Although Miller and Driver[31] tried to make a great distinction between "process" and knowledge, in reality the relationship between scientific skills (science know-how) and scientific knowledge (know-that) is much more subtle and intricate than Miller and Driver[31] tried to believe.

In "Science 5-16: A statement of policy"[29], the Department of Education and Science tried to explain the need of science education for all the pupils. It suggests that knowledge of scientific methods will not only help to develop intellectual talent, but also offer opportunities for careful observation, measurement and planning in everyday life in the future.

In his paper, Gott[33] discussed the nature of some of the skills involved in experimental work and how assessment of these skills can be carried out. He has mentioned 4 major skills, viz. (i) observation, (ii) measurement, (iii) explanation (i.e. interpretation), and (iv) open-ended investigation and problem solving.

Apart from these, I think, pupils will also need to develop the skill of recording and presenting (in table/graph) their observations and measurements of various aspects of an experiment.

For the assessment of observation and measurement, pupils can be provided with a check-list of all possible options. When a pupil records his/her observations or measurements, it may or may not be the truth but it is what the pupil considers to be happening during the experiment. These results can be assessed against some standard results providing a basis for the assessment by Method -P as mentioned earlier. However, for on-the-spot assessment of the performance of the skills/techniques during practical work (assessment by Method -T) several experienced science teachers equipped with check-list, can be used as judges (assessors).

The Bunsen burner is a simple, but versatile apparatus in the chemistry laboratory. A pupil usually starts using it as soon as he enters the 11+ science course and continues to do so throughout his school-life. Experiments involving heating by Bunsen burner are numerous in school science practical courses. In a study, Farmer[34] analysed the RNOC

syllabus for practical work and identified the "key" manipulative tasks and cognitive tasks which were playing important roles, not only in carrying out the laboratory work, but also in the complete understanding of the whole chemistry course. Some of the "key" manipulative skills frequently occurring in the experimental work are:
(1) use and control of a Bunsen burner to heat a solid/liquid,
(2) transfer of liquid (or powdered solid) by pouring,
(3) manipulation of cork/bung + accessories,
(4) operation of direct reading balance and,
(5) electrical connections within electrolysis circuit etc. etc.

It would be worthwhile to subdivide each task into various steps of manipulative skill and judge the pupils' achievement in each step to isolate the area(s) of weakness for which special care can be provided later.

Hubbard and Seddon[35] investigated the performance of pupils (12-14 years old) in heating solids in a test-tube using a Bunsen burner. Attempts were made to assess the abilities of pupils in carrying out the 3 manipulative skills involved. It has been found that almost all the pupils are good in placing the test tube holder at the correct position. However, a large percentage are unable to choose the correct colour of the flame and also to hold the tube at the correct angle. No significant difference in achievement was observed between the sexes and also amongst the 3 year groups (2nd, 3rd and 4th year).

Seddon and Barry[36] also used experiments involving the heating of solids using the Bunsen burner as a part of their investigation. In this study, pupils' ability was judged for carrying out 3 manipulative skills, as mentioned in an earlier work by Hubbard and Seddon[35], in addition to recording of a number of observations. However, the main objective of this work was to compare the achievement of 4 groups of pupils carrying out the same experiment by following 4 different modes of instruction, viz. (1) work-sheet, (2) work-sheet + static diagram, (3) work-sheet + static diagram + audio-tape, and (4) work-sheet + static diagram + video tape with pictures of demonstrations.

It has been found that the pupils who followed only the work-sheet produced the best results. These results have been explained on the basis of the idea that a pupil working with only written instructions can think independently and create his/her own mental imagery needed to plan the practical work. On the other hand when additional materials were provided in the form of another mode of instructions, the pupil's thought has been influenced by the outside idea which may interfere with the independent thinking ability and impede planning and progress. Furthermore with the video film, where lifesize moving pictures of a demonstration were provided, this may have given the pupils a sense of false security, and discouraged them from thinking independently to create a working condition on the bench.

In addition to heating solids, experiments involving the measurement of volume by using a measuring cylinder were also used by Seddon and Barry[36] to confirm their findings regarding the most efficient mode of instruction in a chemistry practical. Various manipulative skills, thought to be needed to measure the volume of a liquid accurately, were checked by the judges. However, the nature of skills, such as if the pupil looks at the meniscus by keeping his/her eye at the correct position or if he/she is holding the cylinder in a correct way, seems to be very difficult to check by a judge without interrupting the pupil's work and within a certain period of time (lesson-time) when 2 groups of 20 pupils have to be assessed.

It would be better to teach these necessary steps in an experimental session during lessons. Furthermore it could be preferable not to assess how a pupil performs the volume measurement, but to check at the end if the volume of the liquid in the cylinder is correct.

Some reasons for the variation in the pupils' ability to retain certain facts taught earlier, have been put forward by Ausubel[37,38]. In his report, he suggested that there are two types of learning --- rote learning, which is easy to forget, and meaningful learning, which is retained by the pupil.

The learning and retention of meaningful materials are primarily affected by their interaction with relevant subsuming concepts already established in the cognitive structure. Materials learned by rote, on the other hand, are discrete and isolated entities which have not been related to any established concepts in the learner's cognitive structure. This is because they are not anchored to an existing identical system, they are much more vulnerable to being forgotten.

Ausubel's assimilation theory of meaningful learning[37,38] and retention has been criticised as "hopelessly vague" by Anderson et. al.[39], who proposed a new theory, called "Schemata theory" which suggests that knowledge is stored in various slots or "placehold" in the learner's mind and this can be instantiated with particular cases.

However, in reply to this criticism, Ausubel[40] wrote the following in a subsequent article:- "It is paradoxically appar~~a~~ent that their theory of schemata, substantively almost identical with my more inclusive theory of anchorage ideas in cognitive structure, which they first seem to approve of, but now stigmatise as "hopelessly vague", and in their actual research, their schemata, unlike mine, do not facilitate meaningful learning".

As the success of a pupil in the assessment of his/her practical work depends not only upon how much he/she can retain the skill/knowledge from the teaching session, but also upon how efficiently they were taught by the teacher, it is perhaps worthwhile to look at the findings of the CLIS project by Driver and Oldham[41]. They have suggested the development of a model for a constructivist teaching sequence. Pupils are thought to engage in activities which encourage them to construct scientific ideas for themselves. This constructivist approach to curriculum development was suggested in response to the findings of an APU project[42-44], which reported that science teaching was didactic and pupils found it difficult to understand certain scientific ideas.

The constructivist view of learning and teaching science, which emphasises that the learner is ultimately responsible for his/her own learning through a sincere effort, is also suggested in another article by Driver and Bell[45].

The process of learning, i.e. the acquisition of the ability to enhance human performance, knowledge structure and conception can be explained both in terms of behavioural psychology and cognitive psychology, although the recent trend is to put more emphasis on the latter than the former. Much research has been done and papers have been published in this field which have been reviewed recently by Shuell[46].

Although the manipulative skills involved in science practical work are psychomotor skills, learning of these skills successfully, retaining them and reproducing/transferring them to a new situation do indeed require a congitive ability of the pupil.

The transfer of skills to a new situation not only depends on how well-organised and broad-based is the training session which leads to meaningful learning, but also depends on whether the trainee can see that the two skills have elements in common, as has been mentioned by Annett and Sparrow[47]. Awareness of this link between two skills in two situations is explained by "the identical elements theory". However, another theory, called "the formal discipline theory", can also be used to explain the transfer of skills. It asserts that two skills depend on a common capacity or potential which can be developed by practising.

In any study of assessment of practical work, its reliability and validity should be looked at before drawing a meaningful conclusion. In this context, the work reported by Eglen and Kempa[48] can be cited. This study questions how concordant (reliable) are teachers' assessments of students' manipulative abilities and how such concordance is affected by different assessment procedures.

The results show that "open-ended" assessment procedures, i.e. those largely based on impression grading without reference to reasonably specific assessment criteria, lead to considerable differences in grades awarded by teachers for the same practical performance.

When assessments are carried out with reference to criteria which relate directly to the performance to be assessed, less divergence in the resulting grades is observed. However, a highly "objectified" assessment mode involving the use of a check-list does not result in full agreement of grades awarded.

In reporting the findings of their study, Geniel and Hofstein[49] mentioned that assessment of practical work using a check-list of well defined criteria (objective mode of assessment) is much more precise and reliable than if the assessment is carried out by observing the overall performance of pupils, and marking by teachers according to their own preference (subjective mode of assessment).

Depending on the nature of the task involved and the apparatus used, assessment of practical work can be carried out either by following a process check-list (on-the-spot test items) and judging on the spot the performance of the manipulative skills involved, or by following a product based "end-check" test item and marking the final outcome of the extended task, i.e. the written reports of the experimental work, containing detailed records of observation, calculation, drawing graphs, making inferences, etc. However, the tendency among the teachers to use the second method is more common as it is often difficult for a teacher to assess a set of 20-25 pupils during a double lesson of 60-70 minutes by following an on-the-spot check list of a reasonable length. The practice of assessing by marking the written report could be considered safe and valid if research could show clearly that there is a strong correlation between the achievement data obtained from two methods of assessment. Even so, assessment of some manipulative skills on the spot during practical sessions is still essential and valuable as science is a practically based subject, and the ability to handle apparatus has its own merit. Cognitive assessment which emphasises factual knowledge recall has been a more widespread activity among the assessors (through examination). Recent moves to "process" in the field of assessment show that assessment of practical skill and abilities, i.e. psychomotor skills, is more problematic. This is because the practically based assessment is more demanding logistically, while the cognitive skill assessment can be done simply on the basis of pencil and paper work. However, all the "processes" (or "steps") involved in the practical work can be divided roughly into two catagories: (i) those that can be done on the spot during practical work and involve the use of equipment, and (ii) those that can be done by intellectual ability or acquired knowledge and involve the use of pencil and paper. The former type of process skill can be called a manipulative or psychomotor skill while the latter type can be called "cognitive skill". In order to perform psychomotor skill correctly, a pupil needs a certain degree of intellectual ability (cognitive skill) to remember what he has learnt earlier by going through a discovery approach and solving the current problems as he progresses through the experiments. Similarly, to be able to perform cognitive skills well, pupils should also do well in manipulative skill performance and learn from that to enhance their knowledge (cognition). So, on the whole, there is no clear cut boundary between these two types of skill, but some degree of overlapping exists.

During their research work on the TAPS project, Bryce et. al.[25] have taken painstaking efforts in developing "process" tests which are content-dependent, i.e. having some form of attachment or relevance to the content of the science which the pupils are studying as a part of their course. As a result of the presence of such content-dependency[31,32] in process tests, pupils will certainly learn much about the scientific content of the topic through the learning of the skill/process involved.

Methods or steps involved in scientific work can be defined[31] as "processes" (some could be manipulative, whereas others could be cognitive) and successful completion of these steps will undoubtedly lead to the knowledge (product) and also learning of the Contents of Science. Scientific processes are not merely pedogogic means/skills/steps/methods used in the field of education, they are much more than that. These processes do have attachment to the

Contents of Science and do help pupils to acquire knowledge and achieve understanding of the subject (i.e. deliver the product).

The term "process" has been viewed by some researchers[31] in two senses of activity: in the pedagogic sense of activity (doing science) and in the scientific/psychological sense of activity (making inferences). Scientific skills (know-how) and scientific knowledge (know-that) are intricately inter-related and the distinction is subtle and elusive[32]. They are interdependent and both form parts of a mental model useful for the learning process.

"Process" in scientific, psychological and pedagogical terms should not be viewed in a limited sense, i.e. that it is free from the contents of science and would not lead to any useful "product" enabling pupils to increase their knowledge and understanding of science.

Many researchers, for example, Miller and Driver[31] consider "Process skill" and "Process" to be the same thing, but Bryce and Robertson[32] make a distinction between these two terms, even through it is a very fine one. By the word "Process", they refer to a skill as such in an unidimensional sense, while "Process skills" seem to mean a sub-set of skills in a particular skill area. For example observing or hypothesising or inferring can be described as "process", while "process skills" are various objectives/steps/activities within the skill area.

However, in this thesis such fine distinction has not been considered in great detail. Simply, all the basic skills in school science practical work are called "Process" skills (they are process of one kind or another and involve some form of skill by the pupil): some involve correct handling of the apparatus and are called manipulative skills, while others involve mental thought process (done by using pencil and paper) and are called Cognitive skills. Each skill may have a number of steps, i.e. a set of skills, which are used as a check-list/marking scheme for on-the-spot assessment during a practical session or when making the written report. Such a report contains aspects of the experimental work which involves cognitive skills and also the final outcome of the whole experiment (i.e. product).

1.4 Purpose of the present study

Seddon and Barry[36] used the reaction of thiosulphate solution with HCl as an experiment for their investigation to find the best mode of instruction out of 4 they used. Each group of pupils following a different mode of instruction was assessed by teachers following a check-list of various manipulative skills. A careful examination of this check-list will reveal that some of them are so simple that every pupil can do well. Some of them will hardly have any influence on the final result, and some of them can be executed so quickly that examiners will find it difficult to assess without interrupting the experiment.

However, having considered these points the check-list can be shortened and a modified one prepared. It has been thought that accurate measurement of the volume of certain liquids and pouring them out completely into the reaction vessel are important manipulative skills capable of influencing the final results. Consequently assessment of these two skills is used to see how effective the pupils are in their practical work. In this study assessments carried out by the judges have not been checked for reliability, nor has the outcome of the assessment been checked for its validity.

At this point, another experiment, e.g. rate of reaction for Mg with dilute HCl, was chosen in which the skill of volume measurement by using a measuring cylinder and pouring it out accurately are involved. It has been thought that these 2 skills are likely to influence the final results of the experiment, i.e. the time needed for the Mg ribbon to dissolve in HCl. In the present study we have assessed the pupils by two methods: (i) Method -T (spot check for both the manipulative skills) and (ii) Method -P (checking the final results, i.e. time for Mg to dissolve, against a standard value). We have also tried to correlate the achievements of a pupil assessed in two different ways. Reliability of assessment carried out by judges marking only one skill as compared to those marking 2 skills, were studied. The results of this section led to the next part of this study.

Almost all the pupils performed the skill of pouring out liquid almost perfectly. However measuring volume was not as uniformly effective and it was thought that there would be room for teaching this skill and to test it after the lapse of a certain period to find out how much they can reproduce that skill. In addition to the volume measurement by measuring cylinder, certain key skills involved in using a Bunsen burner during the heating of solid samples were also taught and assessed after a similar time lapse. This was thought to be important because the use of a Bunsen burner to heat solid samples occurs as a part of many experiments.

This work was a follow-up of the work carried out and reported by Hubbard and Seddon[35] which shows that, in heating solids in a test-tube using a Bunsen burner, pupils can perform some skills very well, but some skills less so. This type of weakness in certain manipulative skills was thought to be responsible for the overall poor performance in using the Bunsen burner by certain pupils.

In the present study some pupils have been given lessons on these skills and tested after a period in order to find out if they can reproduce these skills. Attempts have also been made to compare the performance of these pupils with that of the pupils who were not taught these skills.

Finally, it was thought of interest to extend the investigation into the pupils' ability, not only in retaining, but also in extending and applying a skill taught earlier into a new and slightly more difficult situation.

For this purpose pupils were tested in measuring a certain volume of liquid by using burette and heating a liquid in a test-tube, while they were given lessons on measuring volume by measuring cylinder and heating solid samples in a test tube some time earlier.

Both the methods of assessment used here are based on well defined check-lists and thus can be classified as "objective mode" of assessment which has been thought[48,49] to be more reliable than "subjective mode" of assessment (i.e. assessment based on impression grading by the teachers).

1.5 Scope and limitation involved in designing the experiments for the present study.

In general the experimental materials should be educational --- related and relevant to the course which is being studied by the pupils taking part in the research project. In that way it will be beneficial to them and the time spent by them for our investigation will be worthwhile. Care must be taken so that the experiment is not a repetition of the work that has recently been done during lessons as a part of their class-work.

In preparing a work-sheet, we must remember that each pupil should not only find something to learn, but also find that it lacks repetition. To make the work stimulating and attractive, the apparent objective may be different from the interest of the researchers.

So, we need to incorporate certain elements in the work which interests us as good research material, and to ensure the whole work appears to be interesting, educational and relevant to the tasks pupils have done in the recent past and will do in the future. In other words, they will feel that some form of progress is being made step by step. Furthermore, in addition to the manipulative skills which we wish to study, there should also be sufficient materials involving cognitive skills which motivate more able pupils in doing the work.

To achieve the desired objective, both for the pupils and the researchers, a large number of experiments can be used. However, one must bear in mind the limitation of laboratory facilities in a school. To keep the cost down, chemicals should not be expensive. Furthermore, there should not be any major safety problems in handling the chemicals by the pupils during the experiment.

Regarding the selection of the apparatus, one should be careful to use those items which are available in the school laboratory and likely to be used by pupils during the period of their study in a school.

The amount of apparatus available in a school laboratory may not be adequate for this type of experiment where each pupil will require an individual set. However, this can be overcome by borrowing from the University of East Anglia or a local school.

Laboratory space often influences the number of pupils that can be accommodated in one session. Time for each double-period may be used for each session without causing too much disruption in the daily routine of the pupils. The amount of work to be included in the work-sheet is also carefully planned.

1.6 Statistical methods employed in data analysis

Two important statistical techniques, namely analysis of variance, and test for correlation, were used to analyse the experimental results and understand their statistical significance.

Analysis of variance (ANOVA)

In designing an experiment, the researcher is concerned with the difference between two conditions of a variable. The difference between the subjects' performance under two different conditions of a variable is noted and interpreted. In order to understand the statistical significance of this difference, statistical analysis is carried out. There are various methods for this purpose, but the most suitable method for an experiment depends on the number of variables, number of conditions of a variable, related (matched)/unrelated (different) subject, and the scale (nominal/ordinate/interval etc.) used for data collection for the subjects under different conditions of the variable.

The experiments described in the thesis have two variables, each variable having 2/3 conditions. The subjects were different under each condition and they were obtained by randomising the set of pupils (same age range, similar academic ability and mixed gender). The data (achievement of pupil) were collected on an interval scale. Having considered these aspects, it has become clear that 2-way randomised ANOVA would be appropriate to carry out statistical analysis of the data collected in each of the three experimental design given in Chapter 2, 3 and 4.

Although the frequency distribution of the data was not the completely symmetrical shape of normal distribution in each case, a 2-way ANOVA (a parametric test) could be applied to study the statistical significance of the results, i.e. difference between two conditions of a variable. This is because the subjects' score were gathered numerically on an interval scale and this method gives the opportunity to study the interaction between two variables. Although for a parametric test, homogeneity of variance, i.e. approximately the same variability of score for each experimental condition is a desirable condition, it is not essential as long as there is an equal number of subjects in each experimental condition, and it holds good for each of the experiments described in this thesis.

Furthermore, parametric tests are reasonably robust as far as the above-mentioned conditions (criteria) are concerned, it is unlikely to get a seriously impaired answer about the % probability of obtaining variance ratios. The letter "F" is used to denote the statistics (variance ratio) obtained after completing calculation for an ANOVA and it can be defined as

$$F = \frac{\text{Variability between scores due to independent variable}}{\text{Variability between scores due to other variable}}$$

$$F = \frac{\text{Variance due to source of variance}}{\text{Variance due to error}}$$

According to experimental hypothesis, the variability in score is produced due to a difference in the condition of the independent variable. In order to avoid contributions from other sources, such as intelligence, eagerness to learn, initial knowledge about the subject, gender difference etc., the subjects of the experiment were randomised in groups for studying different experimental conditions. However, one still can claim, giving rise to the Null hypothesis, that the variability is due to other unknown variables than the independent variable manipulated by the researcher.

Analysis of variance gives us the opportunity to obtain a ratio (F) between the variance due to the source of variable, and the variance due to another variance (also known as error factor). If the "F" ratio is high, one can suggest that the variability in score is mainly due to the independent variable and the contribution from other sources is negligible.

Using statistical methods outlined in "The Numbers Game - Statistics for psychology"[50], one can calculate F_A and F_B values for the independent variable (A and B) and also the interaction of the variables F_{AxB} (in case of a 2-way ANOVA).

The significance of F value can then be checked by looking at the critical value of F at the corresponding df (degree of freedom) level given in the data table for critical values of F. If the calculated F value is higher than the critical value for p = (say) 0.01 or 0.05 , then it can be considered as significant at $p < 0.01$ or 0.05 level respectively. This means that the probability of variability in the score due to error factors is less than 1% or 5% respectively and thus the "Null hypothesis" is rejected at this level and the experimental hypothesis is accepted.

A number of 2-way randomised ANOVAs have been used to analyse the experimental results in chapters 2, 3 and 4, as the data are thought to satisfy the necessary conditions described earlier.

Tests for Correlation

Correlation design looks at a possible link/relationship between two sets of data (or score) achieved by a group of subjects under two different experimental conditons (or two different tasks). Here the main objective is to see whether the subjects' performance under two conditions is correlated, either positively, or negatively. As an example, one can cite a pupils' mark in two subjects, e.g. in Physics and Mathematics, obtained by a set of pupils. If there is a positive correlation between the two sets of marks, pupils having the highest mark in physics will have the highest mark in mathematics, while the pupils obtaining the lowest mark in physics will have the lowest mark in mathematics. However, if the correlation is negative, pupils having the high mark in physics will have the low mark in mathematics.

In order to see if there is a correlation, the mathematics score of a pupil is plotted against his/her physics score. This type of plot is known as a scatterogram. It may or may not be possible to visualise easily a clear relationship between two sets of data. Thus a statistical technique can be employed to calculate the correlation value and find if it is statistically significant enough to reject the "Null Hypothesis" (i.e. the apparent relation is due to chance alone) and accept that there is a significant correlation.

It is possible to use either a non-parametric test (Spearman rank correlation coefficient) or a parametric test (Pearson product moment correlation coefficient) to test the significan~~t~~ce of a correlation between two variables. As the data of the experiments reported in this thesis are in an interval scale and generally meet other criteria for a parametric test, Pearson -r values were calculated. The correlation "r" value, is mentioned in terms of a number ranging from -1 (perfect negative correlation), through 0 (no correlation), to +1 (perfect positive correlation).

After calculating the Pearson -r value (for method of calculation, see “The Number Games - Statistics for Psychology”[50]), one needs to look at a statistical table of critical value of “r” to discover the probability of this amount of correlation by the chance factor and thus predict if this ‘γ” value means any significant correlation between the two variables.